

A

USE A SUPERVISED MACHINE LEARNING APPROACH TO INFER A MAPPING FUNCTION F THAT TRANSFORMS THE INPUT FEATURES VECTOR FOR EACH PRODUCT OF THE TRAINING SUBSET OF PRODUCTS TO THE CORRESPONDING AT LEAST ONE PROPERTY FOR EACH PRODUCT OF THE TRAINING SUBSET OF PRODUCTS

~ 160

IDENTIFY BUILDING BLOCK SETS FOR A PLURALITY OF ADDITIONAL PRODUCTS OF THE COMBINATORIAL LIBRARY

~ 170

FORM INPUT FEATURES VECTORS FOR THE PLURALITY OF ADDITIONAL PRODUCTS FROM THE BUILDING BLOCK SETS FOR THE PLURALITY OF ADDITIONAL PRODUCTS

~ 180

TRANSFORM THE INPUT FEATURES VECTORS FOR THE PLURALITY OF ADDITIONAL PRODUCTS USING THE MAPPING FUNCTION F TO OBTAIN AT LEAST ONE ESTIMATE PROPERTY FOR EACH OF THE PLURALITY OF ADDITIONAL PRODUCTS

~ 190

FIG. 1B

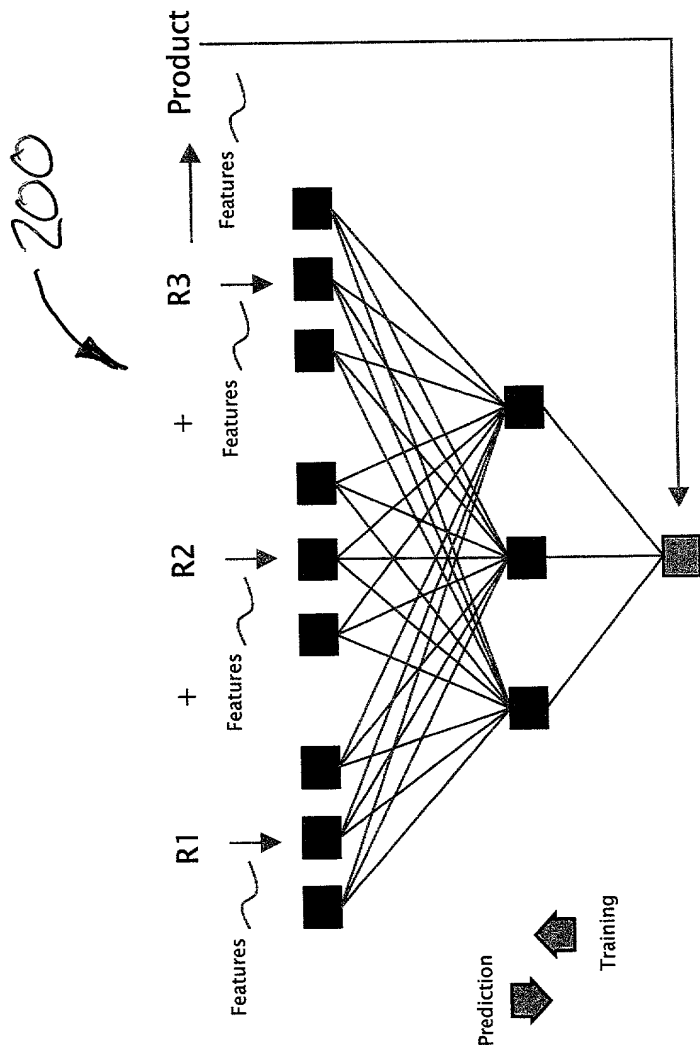


FIG. 2

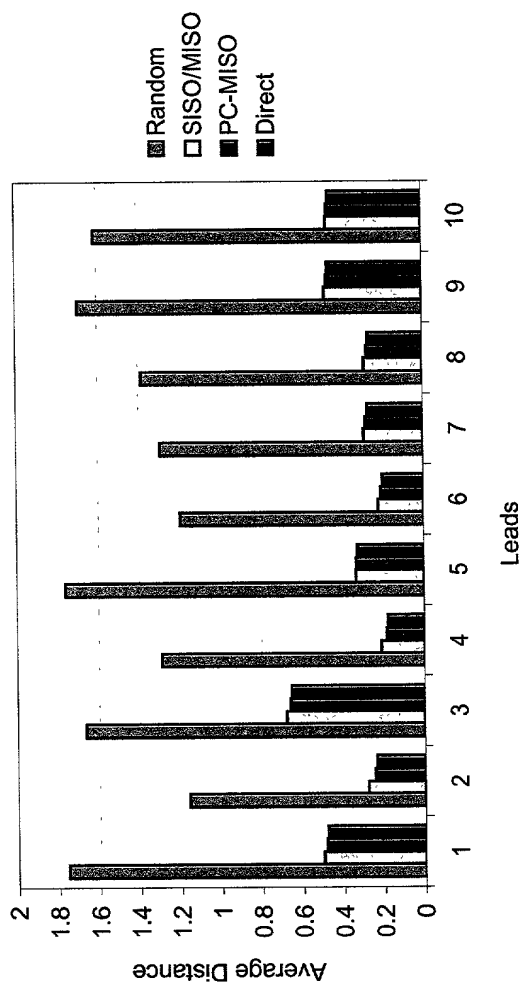
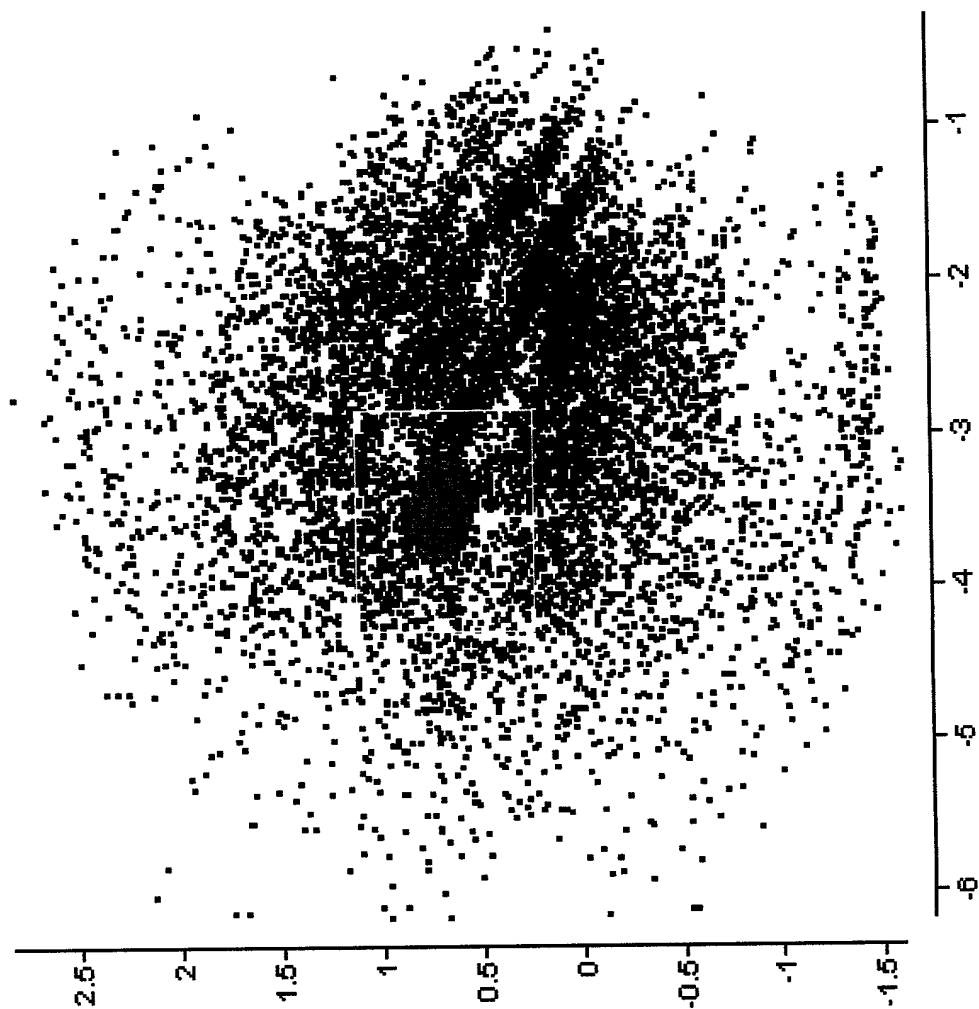


FIG. 3

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	



4A
F16

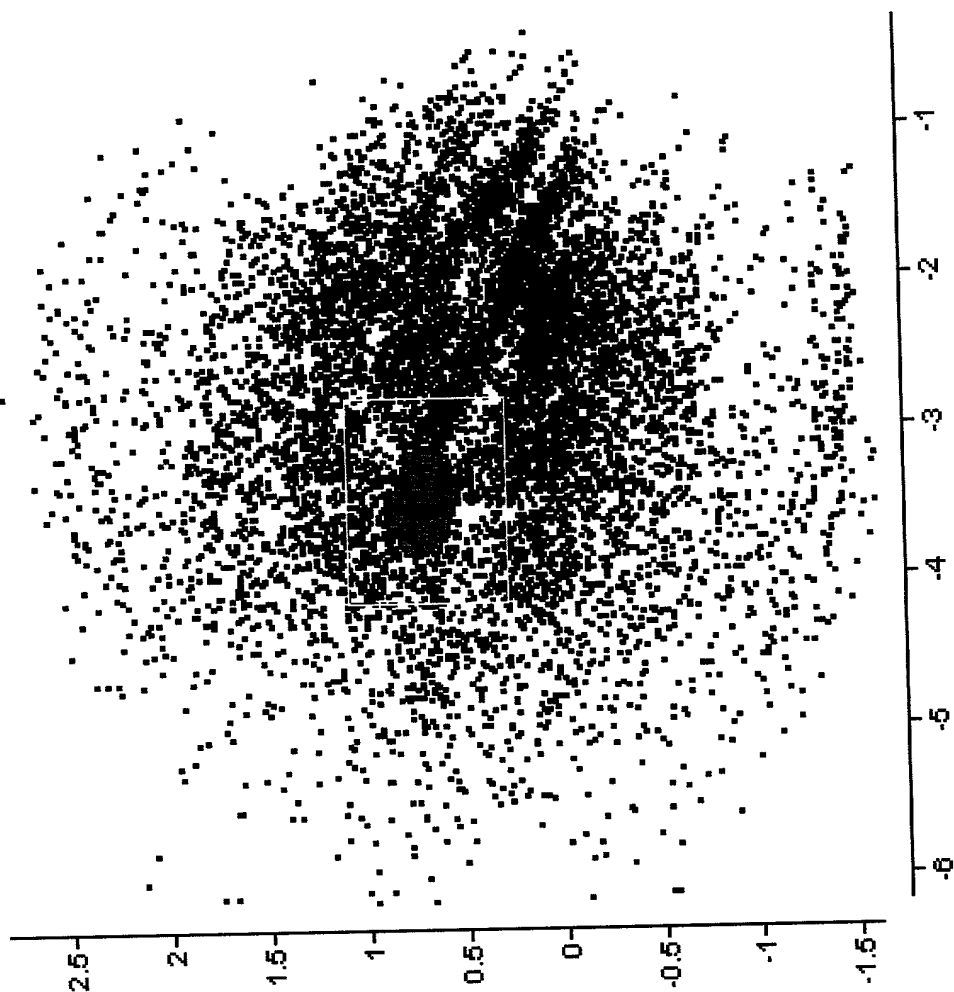
[illegible]

Fig. 4B

1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645
2646
2647
2648
2649
2650
2651
2652
2653
2654
2655
2656
2657
2658
2659
2660
2661
2662
2663
2664
2665
2666
2667
2668
2669
2670
2671
2672
2673
26

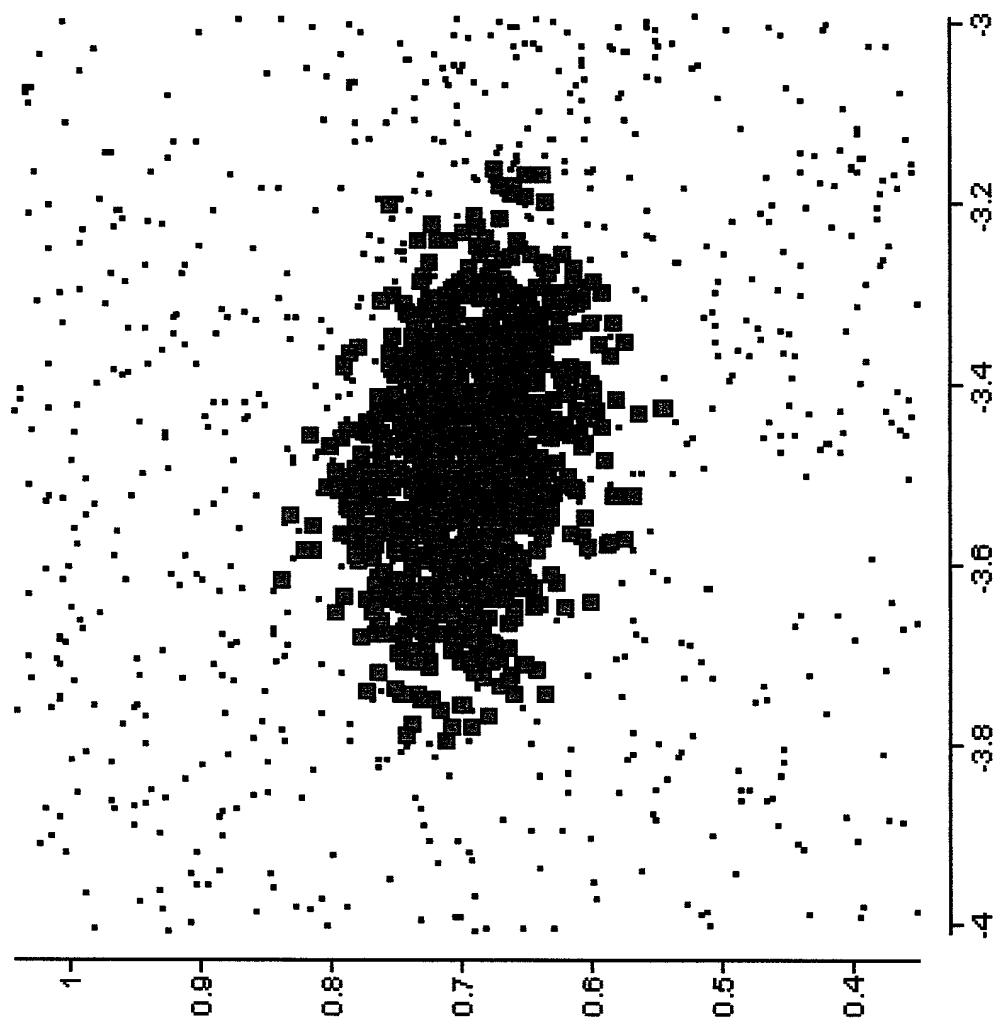
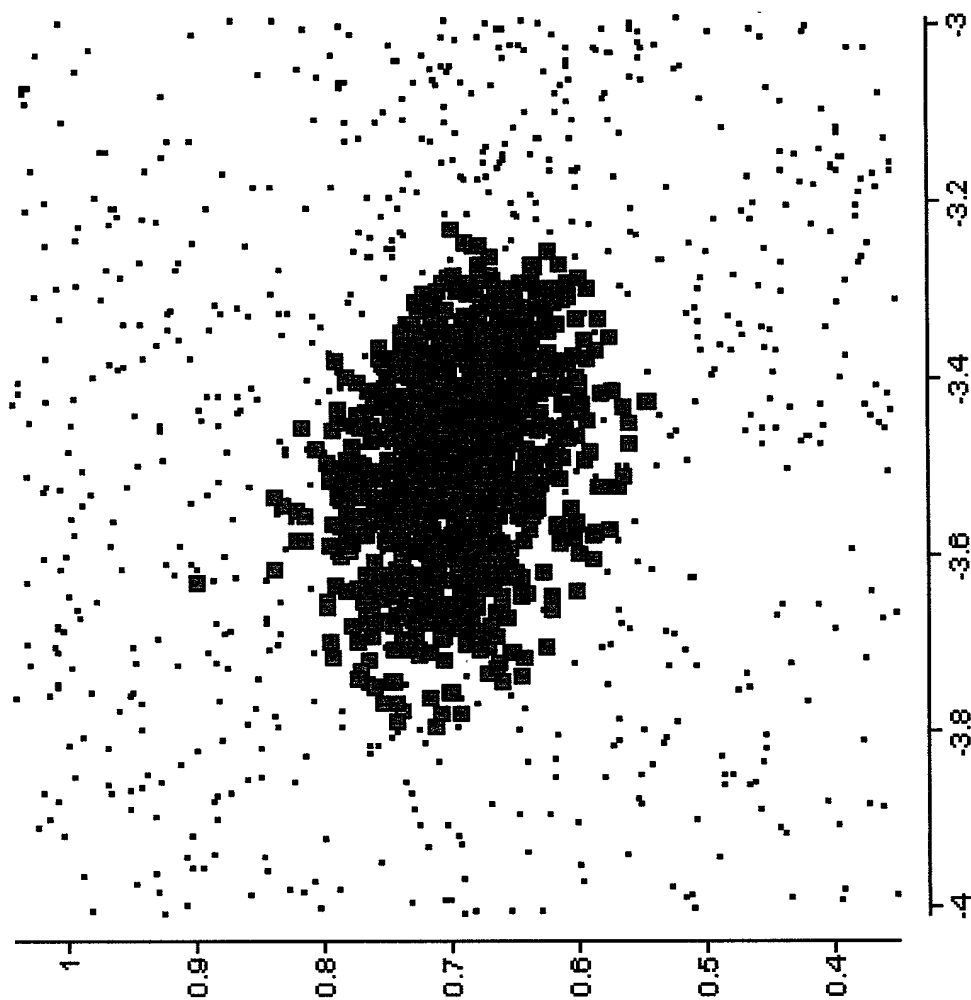
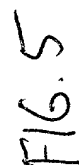


FIG. 4C





FLG. 5



916.6

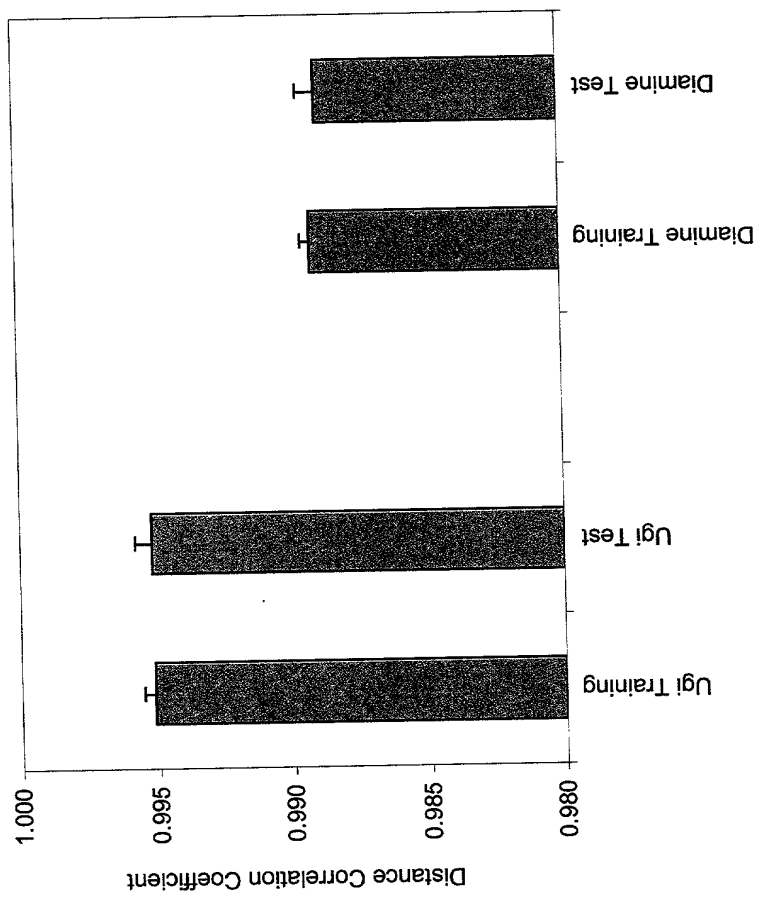
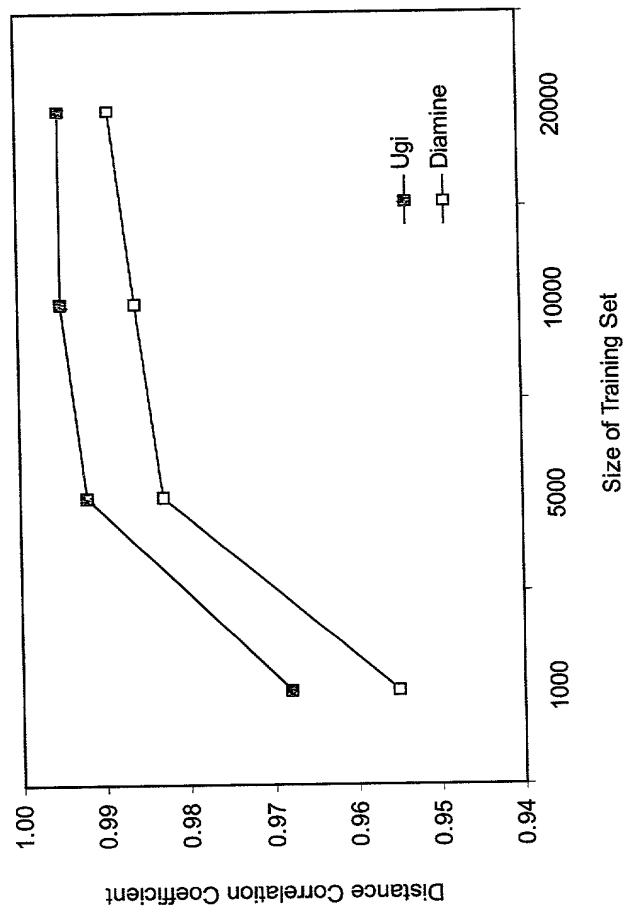


FIG. 7



8/17

Index	Descriptor	SISO Training R ²	SISO Test R ²	MISO Training R ²	MISO Test R ²
1	No. atoms	0.996	0.997		
2	No. bonds	0.995	0.996		
3	No. elements	0.603	0.614	0.822	0.823
4	Molecular weight	0.996	0.997		
5	Chi 0	0.996	0.997		
6	Chi path 1	0.996	0.997		
7	Chi path 2	0.994	0.995		
8	Chi path 3	0.971	0.973		
9	Chi path 4	0.974	0.976		
10	Chi path 5	0.956	0.957		
11	Chi path 6	0.909	0.910		
12	Chi path 7	0.837	0.843	0.943	0.942
13	Chi path 8	0.666	0.673	0.938	0.934
14	Chi path 9	0.563	0.554	0.939	0.936
15	Chi path 10	0.447	0.457	0.950	0.950
16	Chi cluster 3	0.988	0.987		
17	Chi cluster 4	0.993	0.993		
18	Chi path/cluster 4	0.978	0.980		
19	Val chi 0	0.996	0.997		
20	Val chi path 1	0.997	0.998		
21	Val chi path 2	0.996	0.996		
22	Val chi path 3	0.993	0.994		
23	Val chi path 4	0.981	0.982		
24	Val chi path 5	0.952	0.951		
25	Val chi path 6	0.907	0.905		

FIG. 9A

F/G 9B

Index	Descriptor	SISO Training R ²	SISO Test R ²	MISO Training R ²	MISO Test R ²
26	Val chi path 7	0.773	0.775	0.901	0.905
27	Val chi path 8	0.619	0.621	0.890	0.889
28	Val chi path 9	0.349	0.328	0.910	0.910
29	Val chi path 10	0.222	0.201	0.921	0.920
30	Val chi cluster 3	0.994	0.994		
31	Val chi cluster 4	0.993	0.993		
32	Val chi path/cluster 4	0.988	0.989		
33	Chi chain 3	1.000	1.000		
34	Chi chain 4	1.000	1.000		
35	Chi chain 5	0.979	0.978		
36	Chi chain 6	0.995	0.995		
37	Chi chain 7	0.999	0.999		
38	Chi chain 8	1.000	1.000		
39	Chi chain 9	0.999	0.999		
40	Chi chain 10	0.999	0.998		
41	val chi chain 3	1.000	1.000		
42	val chi chain 4	1.000	1.000		
43	val chi chain 5	0.994	0.996		
44	val chi chain 6	0.994	0.995		
45	val chi chain 7	0.998	0.998		
46	val chi chain 8	1.000	1.000		
47	val chi chain 9	0.997	0.998		
48	val chi chain 10	0.986	0.980		
49	subgraph count path 2	0.996	0.997		
50	subgraph count path 3	0.990	0.990		

Index	Descriptor	SISO Training R^2	SISO Test R^2	MISO Training R^2	MISO Test R^2
76	Wiener path no.	0.967	0.965		
77	total Wiener path no.	0.903	0.892		
78	Shannon Index	0.911	0.911		
79	total no. of paths	0.939	0.932		
80	Bonchev-Trinajstić IdW index	0.958	0.955		
81	Bonchev-Trinajstić mean IdW index	0.972	0.972		
82	Bonchev-Trinajstić IdC index	0.979	0.978		
83	Bonchev-Trinajstić mean IdC index	0.793	0.773	0.737	0.759
84	Wiener parity no.	0.988	0.989		
85	Platt F no.	0.996	0.997		
86	Delta partition 1	0.996	0.996		
87	Delta partition 2	0.992	0.992		
88	Delta partition 3	0.997	0.997		
89	Delta partition 4	0.995	0.996		
90	Delta partition 5 ¹	1.000	1.000		
91	Delta partition 6 ¹	1.000	1.000		
92	No. H	0.996	0.997		
93	No. B ¹	1.000	1.000		
94	No. C	0.997	0.998		
95	No. N	0.995	0.995		
96	No. O	0.994	0.993		
97	No. F	0.996	0.996		
98	No. Si ¹	1.000	1.000		
99	No. P	0.999	0.999		

FIG. 9D

Index	Descriptor	SISO Training R^2	SISO Test R^2	MISO Training R^2	MISO Test R^2
100	No. S	0.997	0.999		
101	No. Cl	0.997	0.997		
102	No. Ge ^I	1.000	1.000		
103	No. As ^I	1.000	1.000		
104	No. Se ^I	1.000	1.000		
105	No. Br	1.000	1.000		
106	No. I	1.000	1.000		
107	No. halogens	0.997	0.998		
108	Total topological state 1	0.924	0.918		
109	Total topological state 2	0.947	0.945		
110	Total topological state 3	0.904	0.888		
111	Total topological state 4	0.956	0.956		
112	Total topological state 5	0.852	0.826	0.915	0.907
113	Total topological state 6	0.980	0.980		
114	Total topological state 7	0.832	0.790	0.914	0.898
115	Total topological state 8	0.988	0.988		
116	Total topological state 9	0.913	0.909		
117	Total topological state 10	0.922	0.918		

FIG. 9E

Lead	Random		Direct		SISOMISO		SISOMISO		PC-MISO	
	Similarity		Similarity		Similarity	Identity	Similarity	Identity	Similarity	Identity
1	1.754		0.480		0.501	69%	0.486		0.486	86%
2	1.158		0.238		0.279	56%	0.244		0.244	83%
3	1.664		0.655		0.680	64%	0.660		0.660	84%
4	1.291		0.179		0.213	60%	0.186		0.186	76%
5	1.763		0.327		0.335	82%	0.334		0.334	83%
6	1.196		0.201		0.224	58%	0.209		0.209	75%
7	1.294		0.274		0.291	72%	0.283		0.283	77%
8	1.385		0.268		0.288	73%	0.275		0.275	84%
9	1.694		0.464		0.481	74%	0.470		0.470	86%
10	1.613		0.460		0.470	79%	0.464		0.464	87%

FIG. 10

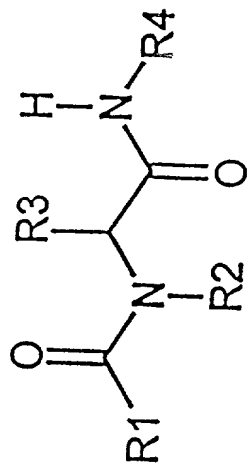
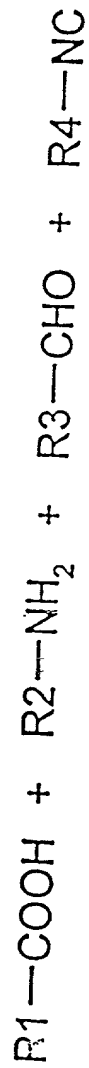


FIG. 11

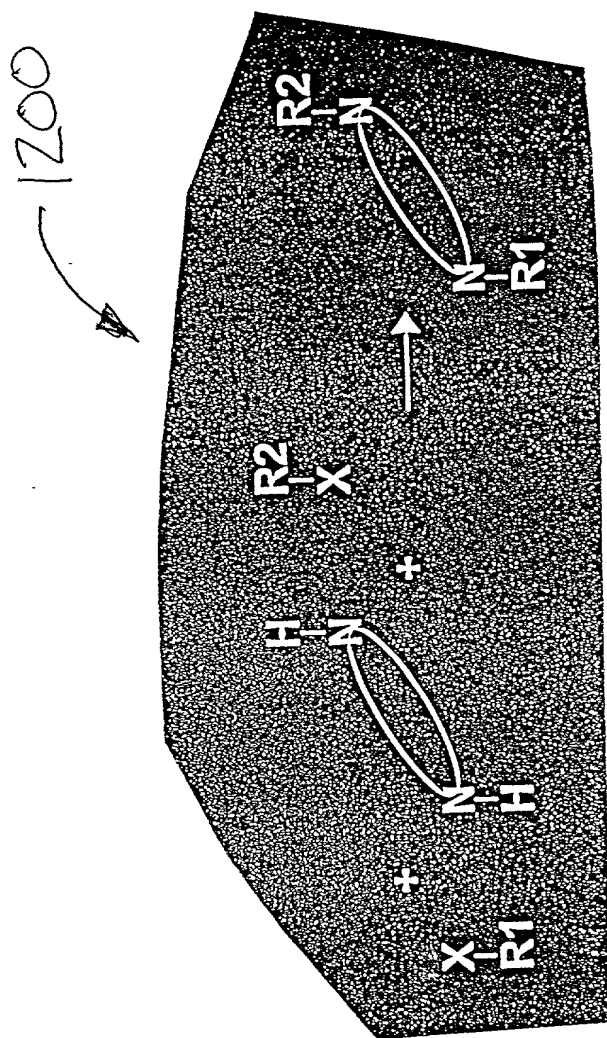


FIG. 12

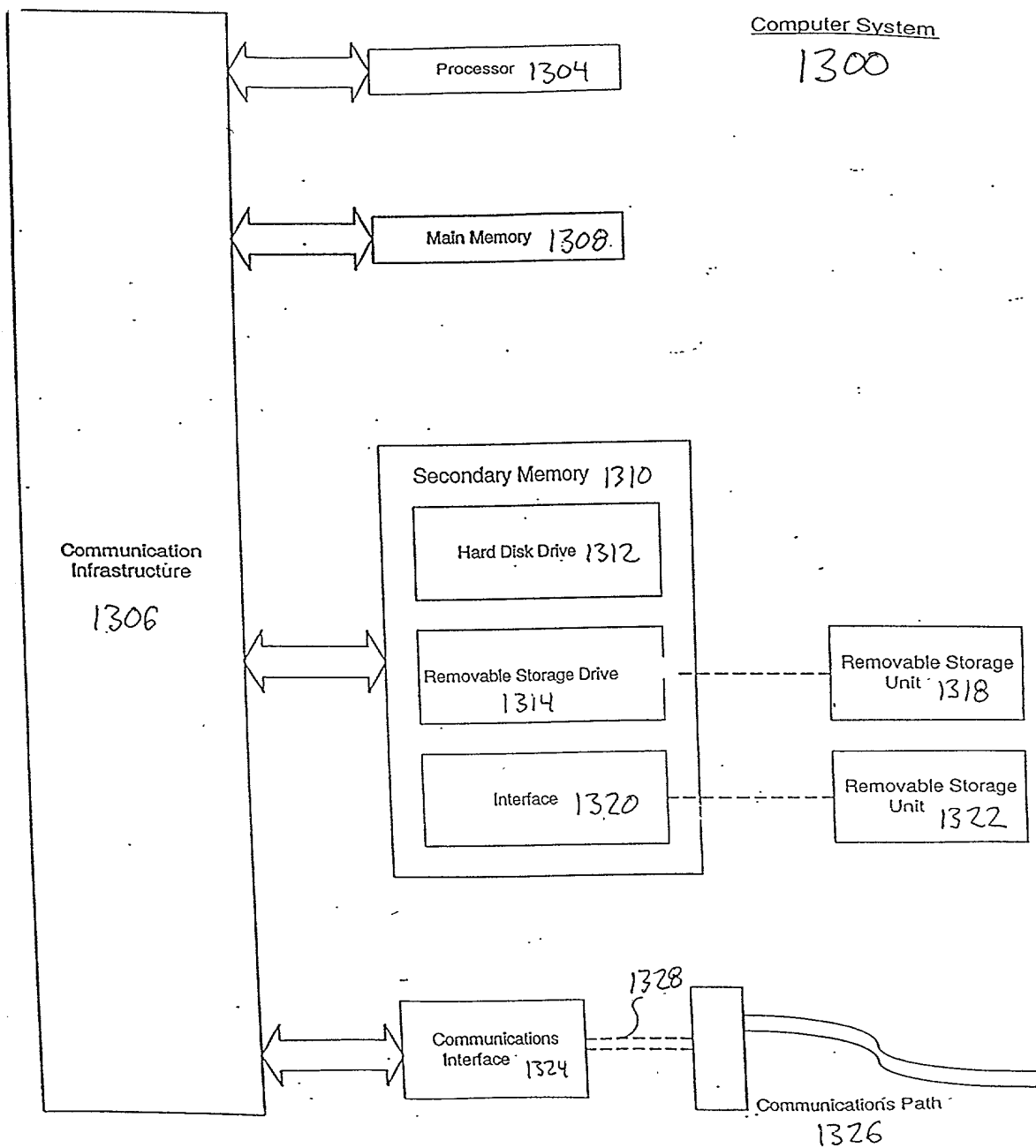


FIG. 13